

## 工作内容

### 简介

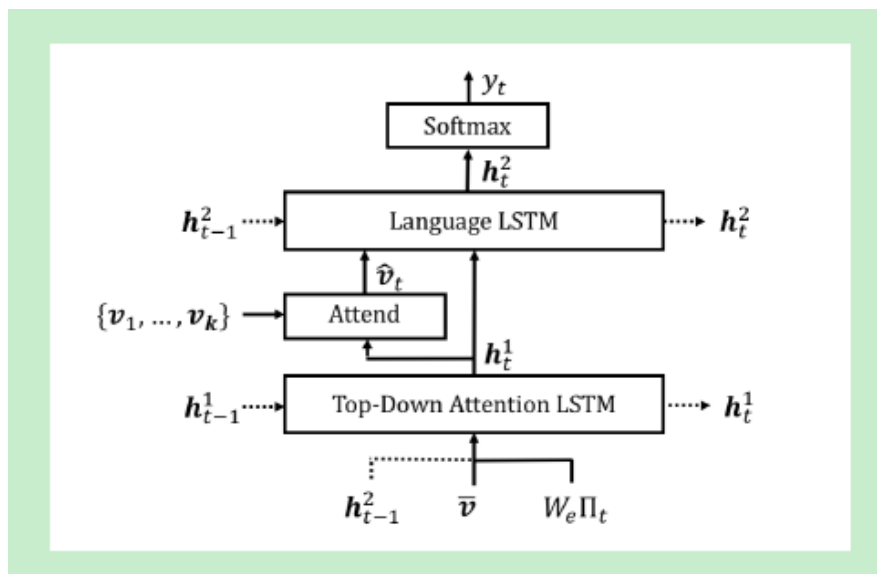
本周工作主要以看论文为主，和夏老师，黄兆嵩博士进行过一些有关未来工作方向的讨论

### 一、《Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering》

**背景：**绝大多数用于图像解释的视觉注意力机制是基于“自上而下”的注意力机制，即将输入图片分成均匀的小格，然后根据生成图像解释词汇的上下文意思，生成“注意力”，而本文则采用 Faster-RCNN 生成候选区域，即自下而上的注意力机制（这一灵感来自于仿生），再结合自上而下的注意力机制，最终取得了 captioning 和 VQA 挑战的冠军

**方法：**输入图像->Faster-RCNN->改进 LSTM（激活函数）->图片解释

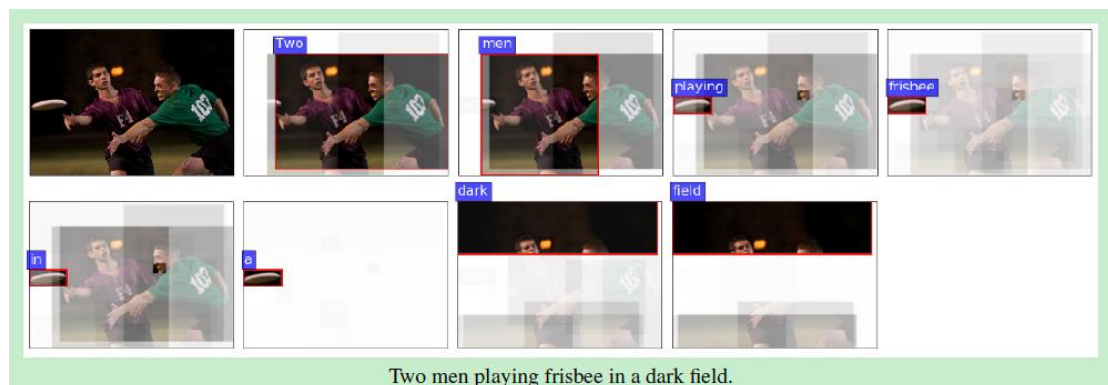
模型结构如下图所示



Captioning 模型

### 贡献：

- 1、首次提出了结合自底向上的注意力机制
- 2、将两种注意力机制进行整合，提出了更易于理解的注意力概念，这一概念和人直观上对注意力的理解一直，本文对注意力的可视化解如下：



借鉴：

- 1、搞清楚了注意力的概念，暂时不知道什么地方会用到，可视化结果的解释的个很有吸引力的领域，但不知道直接把可视化结果转化成文字说明是否有足够的难度

## 二、《DQNViz: A Visual Analytics Approach to Understand Deep Q-Networks》

背景：

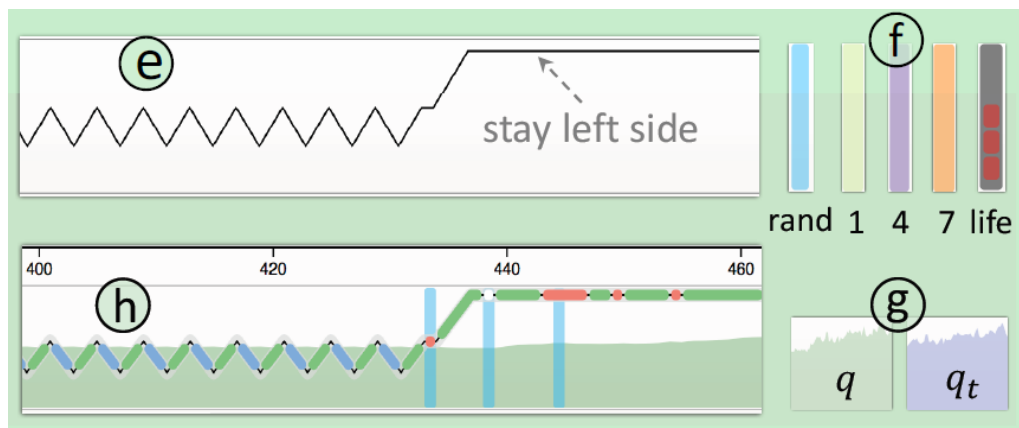
- 1、用可视化的方法去辅助深度学习的训练过程是一个十分有前景的领域
- 2、之前有诸多工作在 vis4dl 上，单少有用可视化解决强化学习的工作
- 3、本文探索性地尝试设计了一个可视分析系统，来辅助解决强化学习的训练问题

方法：

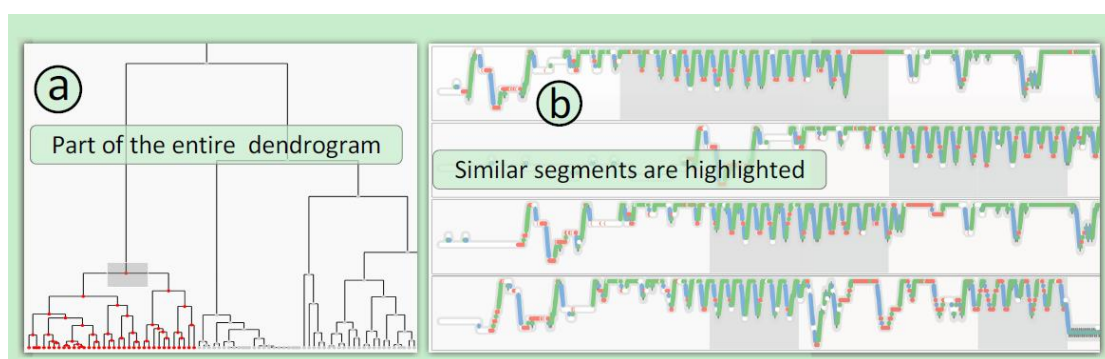
- 1、模型训练的统计数据可视化



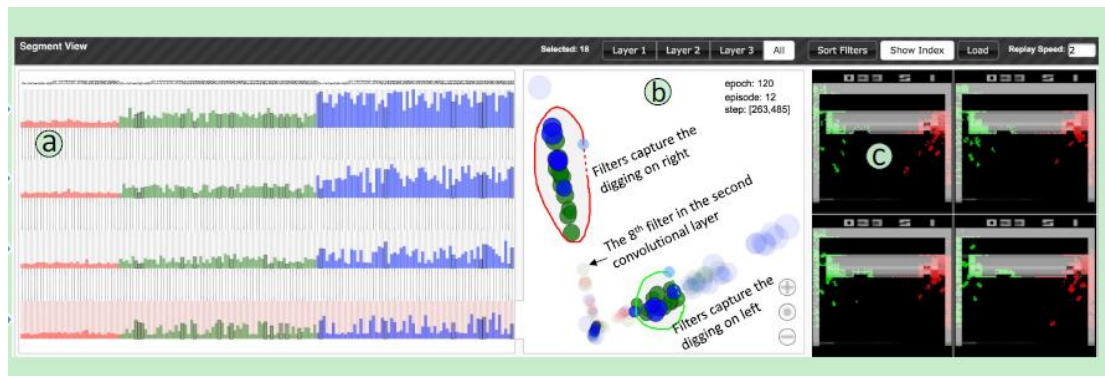
- 2、agent 行为可视化为一个折线图，并用线段的背景色表示 agent 的行为



- 3、将大段的时间片分成小段，利用层次聚类发现有用模式



4、利用神经网络的最大化激活方法对 DQN 的内部结构进行可视化，并对有用结果绘制柱形图，显著图，以及将显著图利用 PCA 降维方法绘制在 2D 坐标系上以发现模式



**贡献：**

- 1、提出了一种综合性的强化学习可视分析系统
- 2、提出了一系列可用于可视分析的强化学习指标
- 3、是一种针对时序数据的可视分析系统，易于发现各种有用的模式

**借鉴：**

- 1、为可视分析技术解决深度学习问题提供了思路
- 2、不足：设计的可视分析系统针对的问题于狭小，不利于推过到其它问题

### 三、《A Perception-Driven Approach to Supervised Dimensionality Reduction for Visualization》

**背景：**

- 1、现有的降维方法没有考虑降维之后将不同类别之间的数据进行区分开来
- 2、现有的类别分离度量没有考虑到类内密度这个特性

**方法：**

- 1、针对 DSC 和 KNNG 分离度量指标存在的不够准确问题，提出了新的计算公式进行了改进
- 2、针对之前方法计算开销大，提出了采用模拟退火方法，取得最优化投影矩阵

**贡献：**

- 1、针对前人工作推导出了一个数学公式，用以衡量人对数据分离度的感知，并用 user study 证明其符合人的感知

$$s(y_i) = \frac{b(y_i) - a(y_i)}{\max\{a(y_i), b(y_i)\}}, \quad (9)$$

- 2、将该度量用在相应的降维方法，可取得类内聚集，类间分离的散点图
- 3、将本方法扩展到了不平衡数据和未标注数据上

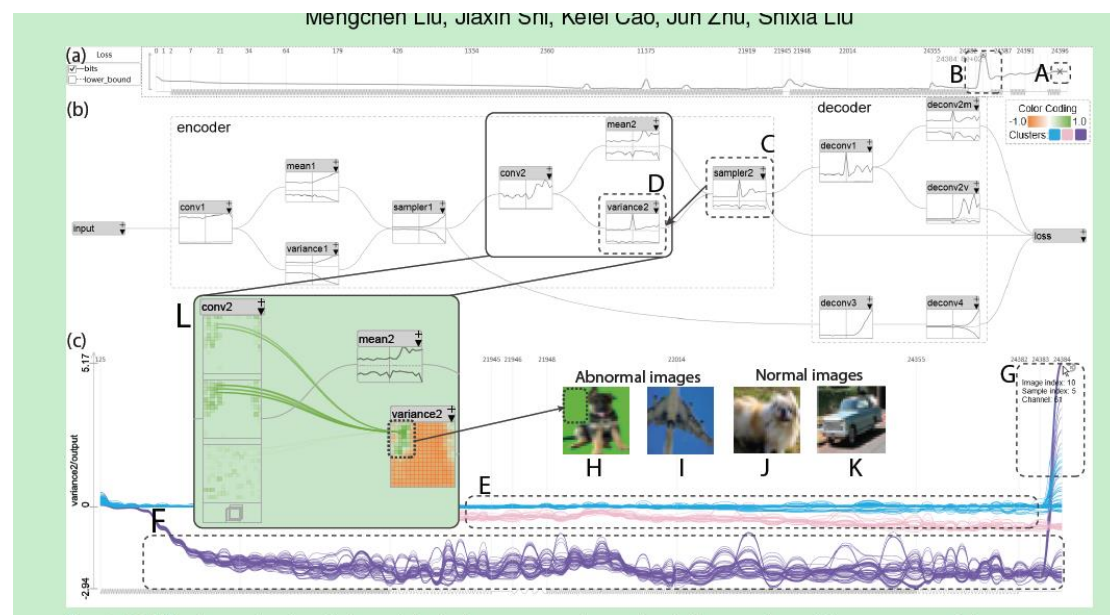
**借鉴：**

- 1、优点：方法具有一定的创意，用简单的数学公式准确刻画人对类别之间的分离度的感知，且具有一定的实际应用价值

- 2、缺点：提出的方法专业性侧重于在降维中保持类别之间的分离特征，可能会丢失其它的特征。且该过程为了可视化效果，将类间的部分特征突出了，即对原始的数据可能会造成一定的失真(也许)

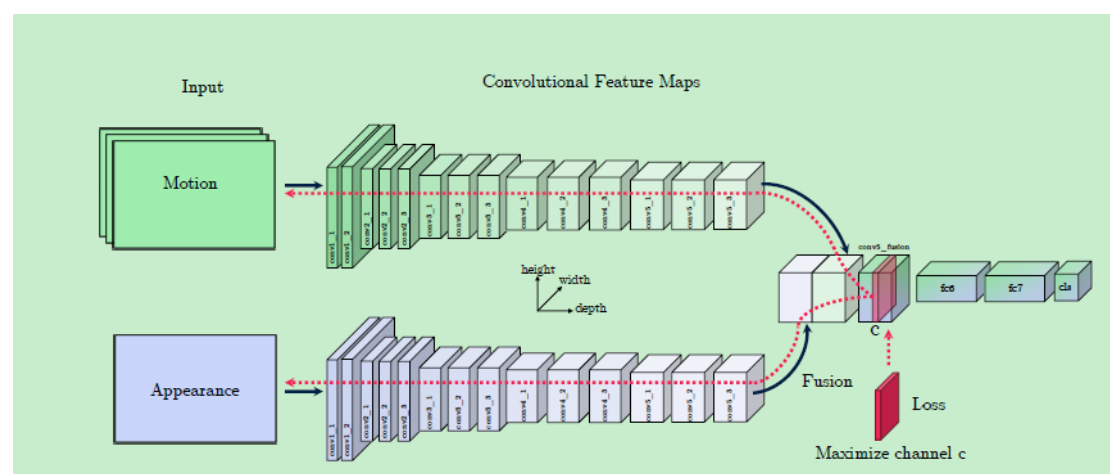
#### 四、《Analyzing the Training Processes of Deep Generative Models》

**简介：**由于更好的 DGM 效果需要 DGM 专家对训练的结果有更深入的理解，本文设计了一种可视分析工具，辅助 DGM 网络专家理解和诊断 DGM，主要采用了蓝噪声采样法。界面如下：

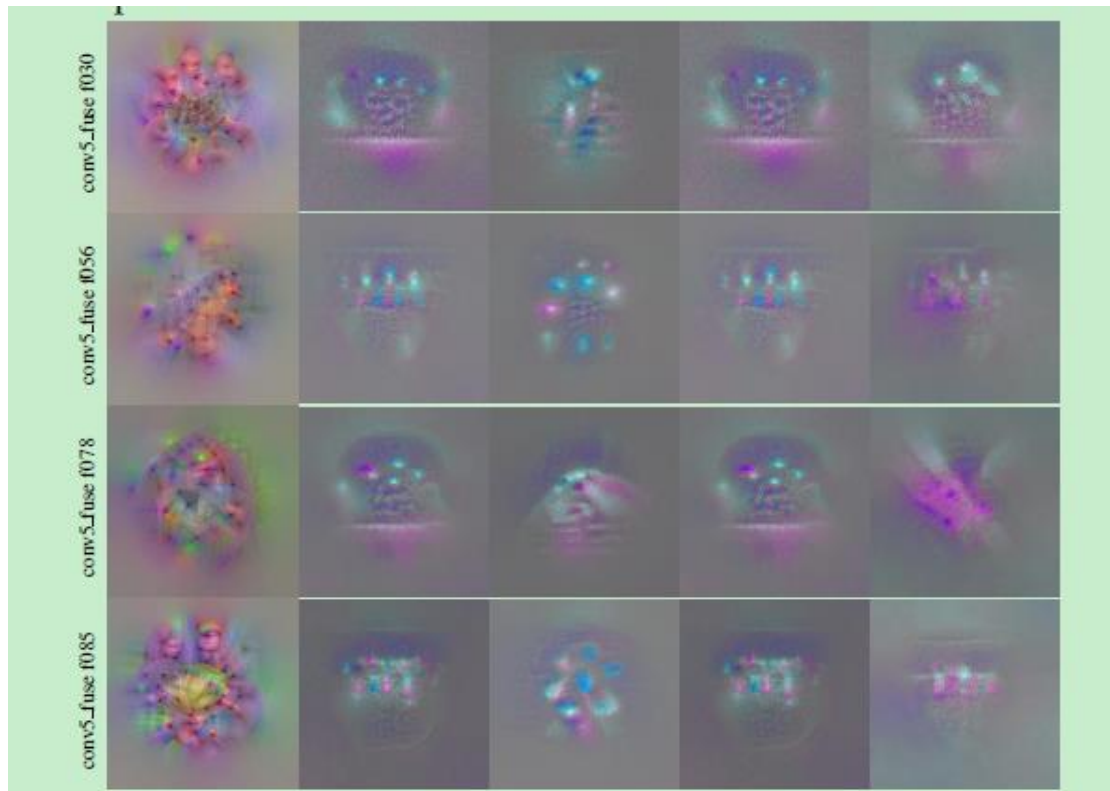


## 五、《What have we learned from deep representations for action recognition?》

简介：本文结合激活最大化方法和生成对抗方法探究深度神经网络的表达，并将该可视化结果的扩展到了时空域上，本文主要是一篇理解深度神经网络的论文，其使用的生成对抗网络如下：



可视化结果如下：



思考：我们能否采用生成对抗网络生成可视化结果？(又一个可能的方法—dl4vis)

## 六、工作时间

周一至周日：9:00-21:30